

ON THIN LOCAL SETS OF THE GAUSSIAN FREE FIELD

AVELIO SEPÚLVEDA

ABSTRACT. We study how small a local set of the Gaussian free field (GFF) in dimension d has to be in order to ensure that this set is thin, which loosely speaking means that it captures no GFF mass on itself. We provide a criteria on the size of the local set for this to happen, and on the other hand, we show that this criteria is sharp by constructing small local sets that are not thin.

1. INTRODUCTION

The Gaussian Free Field (GFF) is the natural analogue of Brownian motion when the time-set is replaced by a d -dimensional open domain D . The GFF is a fundamental object in probability and statistical physics. In two dimensions its geometry is closely related to many other key objects such as Stochastic Loewner Evolutions [7, 14, 23], Conformal Loop Ensembles [5, 13], Liouville Quantum Gravity [2, 8], Quantum Loewner Evolutions [15, 16] and Loop Soups [11, 12, 25]; note that the relation to loop-soups is in fact not restricted to the two-dimensional GFF.

Unlike Brownian motion, when $d \geq 2$, the GFF is not a continuous function; it is only defined as a random generalized function from D into \mathbb{R} . However, the GFF shares many of the Brownian motion's properties and in particular its Markov property; loosely speaking, the spatial Markov property of the GFF is that for any deterministic closed set A the distribution of the GFF in the complement of A is equal to the sum of the harmonic extension of the values of the GFF on ∂A with an independent GFF in $D \setminus A$. Just as in the one-dimensional case, this Markov property can be upgraded into a strong Markov property, where the above decomposition holds also for some random sets A . Such multivariate Markov properties were first studied in the 70s and 80s [21], and recently reinterpreted and applied in the two-dimensional imaginary geometry framework [14, 23]. These sets, called local sets in [23, 14], play roughly the same role, in the higher-dimensional setting, as stopping times; more precisely, the local set A is the analogue of the interval $[0, \tau]$ when τ is a one-dimensional stopping time. The notion of local sets makes sense and is natural for the GFF in any dimension, even if so far it has only been used when $d = 2$.

One way to formally describe local sets is to say that there exists a coupling (Γ, A, Γ_A) where Γ is a GFF in D , A is a random closed set and Γ_A is a random field with the following properties:

- Conditionally on (A, Γ_A) , the distribution of $\Gamma - \Gamma_A$ is a GFF in $D \setminus A$.
- For every deterministic open set O , on the event where O and A are disjoint, the restriction of Γ_A to O is a harmonic function in O . More precisely, there exists a random harmonic function h_A in $D \setminus A$ such that for all smooth function f , $(\Gamma_A, f) = \int_{D \setminus A} h_A(x) f(x) dx$ on the event where the support of f is contained in $D \setminus A$.

The field Γ_A can be loosely speaking understood as being equal to the field Γ “within A ” and as being equal in $D \setminus A$ to the harmonic extension h_A of the values of the field on ∂A .

In the present paper, we investigate how small a local set has to be (for instance in terms of its fractal dimension) in order to ensure that, loosely speaking, Γ_A consists only of its harmonic part h_A , i.e, it carries no mass of the GFF on itself – we call such sets thin local sets: In the very special case where the harmonic function h_A is a.s. integrable on $D \setminus A$ (this for instance happens

This work was supported by the SNF grant #155922. The author is part of the NCCR Swissmap.

for the bounded-type thin local sets studied in [5] where h_A is bounded), being thin means that for any compactly supported smooth function f , (Γ_A, f) is almost surely equal to $\int_{D \setminus A} h_A(z) f(z) dz$, even when the support of f intersects A . In the general case, where the local set is such that the function h_A is not integrable on $D \setminus A$, which should be thought of as the generic case (mind that h_A oscillates wildly when it approaches A , especially in higher dimensions – this is already the case when A is a deterministic non-polar set), there are various possible definitions that we will discuss, but we can sum it up in saying that thin local sets are the local sets for which for all given reasonable procedure to make sense of the not-absolutely-converging integral $\int f(z) h_A(z) dz$ turns out to be almost surely equal to (Γ_A, f) .

Thin local sets are typically small. For instance, a deterministic set is a thin local set if and only if it is of zero Lebesgue measure. But, as we shall see, when $d \geq 2$ there exist many (random) non-thin local sets that have zero Lebesgue measure. In some sense, this is due that one can explore GFF values in such a way to capture large values of the GFF while keeping the explored domain local and fairly small.

Let us briefly present our main results first when $d \in \{3, 4\}$, then $d = 2$ and then $d \geq 5$. For $d = 3, 4$, we have:

- (1)_d If A is a local set of the GFF with upper Minkowski dimension that is almost surely smaller than $(d/2) + 1$, then it is a thin local set.
- (2)_d There exist local sets of the GFF with upper Minkowski dimension that is almost surely not larger than $(d/2) + 1$ that are not thin local sets.

In other words, the dimensions $5/2$ and 3 play an important role for the size of local sets of the GFF in respective dimensions $d = 3$ and $d = 4$.

These statements also hold in the two-dimensional case, but the second one is rather void because $1 + (d/2) = 2$, so that one can just take A to be the entire domain \overline{D} , which is clearly not thin. We derive the following more refined result when $d = 2$:

- (1)₂ If A is a local set of the two-dimensional GFF such that for some positive δ , the expected value of the area of the ε -neighborhood of A decays almost surely like $O(|\log \varepsilon|^{-1/2-\delta})$, then it is a thin local set.
- (2)₂ There exist local sets of the two-dimensional GFF for which the expected value of the area of their ε -neighborhood decays almost surely like $O(|\log \varepsilon|^{-1/2})$ and that are not thin local sets.

When $d \geq 5$, another phenomenon related to the dimension of polar sets enters into the game. We shall prove that when $d \geq 5$,

- (1)_d If A is a local set of a d -dimensional GFF and has upper Minkowski dimension smaller than $\max\{d - 2, 1 + (d/2)\}$, then it is thin.
- (2)_d There exist local sets of the GFF with upper Minkowski dimension almost surely equal to $d - 1$ that are not thin local sets

We believe that one can replace $d - 1$ by $d - 2$ in the statement (2)_d. The threshold $(d/2) + 1$ would then be valid up to $d = 6$, and for $d > 6$, it should therefore be $d - 2$.

Our proofs of statements of the type (1)_d (i.e. “when the local set is small enough, then it is necessarily thin”) are based on rather direct moment estimates: When $d \neq 2$, a first moment computation combined with a Borel-Cantelli argument suffices, and when $d = 2$, we use a slightly more refined second moment computation.

It is somewhat more challenging to prove (2)_d, i.e. to construct well-chosen “fairly small” local sets and to prove that they are not thin, and this is arguably the main contribution of the present paper. It is worthwhile noticing that in two-dimensions, it is possible to use the nested version of the Miller-Sheffield GFF-CLE₄ coupling to construct such a small yet non-thin local set [3], but when $d \geq 3$ other ideas are needed. Our strategy consists in relating a particular exploration of the

GFF with a branching Brownian motion. This idea is reminiscent of the one that was for instance used in the two-dimensional case in [6] to study the maximum of the discrete GFF. The constructed set may also be interpreted as a local set approximation of perfect thick points (in the sense of [10], Section 3.2).

The structure of the paper is the following: we first very briefly recall some basic properties of the continuous GFF, its local sets and we give a possible definition of thin local sets. Then, we use this definition to construct examples of local sets that prove the statements $(2)_d$. Thereafter, we prove the statements $(1)_d$ and conclude with some comments about the definitions of thin local sets.

Acknowledgements. First, I wish to thank my advisor Wendelin Werner for having proposed me the problem, for all the inspiring discussions and specially for all the many times that he read the many manuscripts proposing each time many changes that made the paper (much more) readable. I also wish to thank Juhan Aru, for the inspiring and useful conversations and the (many) times he read this manuscript helping it to improve it every time. Finally, I wish to thank Ron Rosenthal for the stoicism he showed when reading the very first version of this manuscript and commenting it.

2. PRELIMINARIES

2.1. GFF and scaling. Introductions and basic results about the GFF can be found in [1, 4, 23, 26, 28]. While the presentations in those references is in the two-dimensional setting, they can be extended without any difficulty to higher dimensions. Let us briefly remind some basic facts:

Throughout this paper, we will use the function ϕ_d defined on $\mathbb{R}^d \setminus \{0\}$ by $\phi_d(x) = (1/2\pi) \times \log(1/\|x\|)$ when $d = 2$ and by $\phi_d(x) = 1/(c_d\|x\|^{d-2})$ when $d \geq 3$, where c_d denote the $d - 1$ -dimensional surface of the unit sphere in \mathbb{R}^d .

Suppose that D is d -dimensional open domain with non-polar boundary (this boundary can be empty if $d \geq 3$), and consider the Green's function with Dirichlet boundary condition in D to be the unique function from $D \times D \setminus \{(x, x) : x \in D\}$ to \mathbb{R}_+ that is harmonic in both variables, and such that for all given x in D , $G_D(x, y) \rightarrow 0$ as $y \rightarrow \partial D$ and $G_D(x, y) \sim \phi_d(x - y)$ as $y \rightarrow x$. Recall that when $D \subset \tilde{D}$, then $G_D(x, y) \leq G_{\tilde{D}}(x, y)$.

We can then define the space $\mathcal{H}^{-1}(D)$ of functions on D , such that

$$\iint_{D \times D} f(x)G_D(x, y)f(y)dxdy < \infty.$$

The GFF in D with zero boundary conditions is defined to be the centered Gaussian process $((\Gamma, f), f \in \mathcal{H}^{-1}(D))$ with covariance function

$$\mathbb{E}[(\Gamma, f)(\Gamma, g)] = \iint_{D \times D} f(x)G_D(x, y)g(y)dxdy.$$

It is well-known that this process exists, and that it is possible to find a version of the GFF such that almost surely, for all $\varepsilon > 0$, Γ can be viewed as an element of the space $\mathcal{H}^{1/2-d/4-\varepsilon}$, the dual under the \mathcal{L}^2 product of the Sobolev space $\mathcal{H}^{d/4-1/2+\varepsilon}$ (see for instance Section 2.3 of [26]).

The definition of the GFF immediately implies its scaling properties. If we define the domain $z_0 + rD := \{z_0 + rz : z \in D\}$, then

$$(2.1) \quad G_{z_0+rD}(z_0 + rx, z_0 + ry) = r^{2-d}G_D(x, y)$$

(in two dimensions, a stronger result holds, as the Green's function is conformally invariant), which yields the corresponding scaling properties for the GFF.

2.2. Local sets. We first very briefly review the definitions of local sets and some of their properties that are relevant for our purposes.

Denote the family of all closed subsets of D by $\mathcal{C}(D)$. Let Γ be a GFF in D and $C \in \mathcal{C}(D)$, one can decompose Γ into the sum of two independent processes Γ_C and Γ^C where almost surely, Γ_C restricted to $D \setminus C$ is a harmonic function, and where Γ^C is a GFF in $D \setminus C$ (this property is usually referred to as the spatial Markov property of the GFF). One can note that Γ_C and Γ^C are Gaussian processes that are also generalized functions, with respective covariance given by the Green's functions $G_D - G_{D \setminus C}$ and $G_{D \setminus C}$.

Let $(\mathcal{F}_C)_{C \in \mathcal{C}(D)}$ be a complete outside-continuous filtration indexed by $\mathcal{C}(D)$ i.e. $C \mapsto \mathcal{F}_C$ is non-decreasing, the σ -fields \mathcal{F}_C are all complete with respect to the probability measure that we are working with, and for any decreasing sequence (C_n) , one has $\mathcal{F}(\cap C_n) = \cap \mathcal{F}(C_n)$. We say that the GFF Γ is adapted with respect to this filtration if for all C , Γ_C is \mathcal{F}_C measurable while Γ^C is independent of \mathcal{F}_C . We also say that a random set A is a local set in the filtration (\mathcal{F}_C) if for all $C \in \mathcal{C}(D)$, the event $\{A \subset C\}$ is in \mathcal{F}_C . The filtration generated by a GFF Γ (or the “natural filtration” of Γ) is the smallest one for which each Γ_C is \mathcal{F}_C -measurable.

Let us list a couple of simple facts about local sets, whose properties are immediate consequences of the definition (see [4]):

- a) If A and B are local with respect to the filtration (\mathcal{F}_C) , then $A \cup B$ is also local.
- b) If (A_n) is a family of local sets with respect to the filtration (\mathcal{F}_C) , then $\cap_n (\overline{\cup_{m \geq n} A_m})$ is also a local in the same filtration.
- c) If A is a local set and Γ is a GFF adapted to \mathcal{F} , then there exists a process Γ_A , such that it is a.s. harmonic in $D \setminus A$, and that conditionally on (A, Γ_A) , $\Gamma^A := \Gamma - \Gamma_A$ is a GFF in $D \setminus A$.

In the literature, having a coupling (A, Γ) satisfying c) is usually used as a definition of local set (see for instance [23]). This property is equivalent to the existence of a filtration under which A is a local set and Γ is a GFF (see [4]). The definition of local via filtration will be handy to show that the examples that we construct are indeed local sets.

Note that we can represent the restriction of Γ_A to $D \setminus A$ as a harmonic function h_A in $D \setminus A$, i.e., there exists a harmonic function h_A in the random domain $D \setminus A$ such that for all smooth function f with compact support in D , $(\Gamma_A, f) = \int h_A(z) f(z) dz$ almost surely on the event where the support of f is contained in $D \setminus A$.

Additionally, it holds that when A and B are local sets, a.s. for all z such that the connected component of $D \setminus A$ containing z is equal to the connected component of $D \setminus B$ containing z we have that $h_A(z) = h_B(z)$ (see [4, 26, 28]).

Let us already point out that local sets have to be big enough in order to actually provide any information about the GFF:

Lemma 2.1. *Let Γ be a GFF on a domain D and A a local set. Then, $\Gamma_A = 0$ almost surely if only if A is almost surely polar for Brownian motion on D .*

Proof. Note that A is polar if and only if $G^D = G^{D \setminus A}$. Then for all smooth function f with bounded support,

$$\mathbb{E}[(\Gamma_A, f)^2] = \mathbb{E}[(\Gamma, f)^2] - \mathbb{E}[(\Gamma^A, f)^2],$$

Given that $G^{D \setminus A} \leq G^D$, we see that A is polar if and only if the right hand side is equal to 0 for all such f . \square

Recall that Kakutani's Theorem (Theorem 8.2 in [18]) shows that being polar or not is in fact just a condition on the decay on the volume of small neighborhoods of A . In particular, we see that when $d \geq 3$, any local set with Minkowski dimension smaller than $d - 2$ is polar for the BM, and it is therefore a local set with $\Gamma_A = 0$.

2.3. A first possible definition of thin local sets. We will discuss in more detail various possible definitions of thin local sets and questions related to those definitions in the last section of the paper, but at this point, let us already give here a definition based on dyadic approximations. We will work with this definition in the coming sections.

Suppose that D is a fixed open domain in \mathbb{R}^d for $d \geq 2$. For simplicity, we assume that D is a bounded set. For any $n \geq 0$, say that s is an open dyadic hyper-cube of side-length 2^{-n} (or just 2^{-n} dyadic hypercubes) if it is a translate of $(0, 2^{-n})^d$ by some element in $(2^{-n}\mathbb{Z})^d$. We call \mathcal{S}_n the set of all non-empty intersections of open 2^{-n} -dyadic hypercubes with D and \mathcal{T}_n the set of faces of elements of \mathcal{S}_n . If A is a closed set we define A_n to be the closure of the union of elements of $\mathcal{S}_n \cup \mathcal{T}_n$ intersecting A .

With this definition, we can note that if A is a local set for the GFF Γ in D , A_n is also a local set. Note that $A_n \searrow A$ and for each $n \in \mathbb{N}$, A_n can take only finitely many possible values. This second fact makes it possible to define, for each smooth bounded function f in D , random variables $(\Gamma_A, f1_{D \setminus A_n})$ and $(\Gamma_A, f1_{D \setminus A_n})$. Indeed, one can simultaneously define $(\Gamma_A, f1_u)$ for any possible value u of $D \setminus A_n$, and then $(\Gamma_A, f1_{D \setminus A_n}) = \sum_n (\Gamma_A, f1_u) 1_{\{D \setminus A_n = u\}}$.

Definition 2.2. *We now say that the local set A is thin if for any smooth bounded function f in D , the sequence of random variables $(\Gamma_A, f1_{D \setminus A_n})$ converges in probability to (Γ_A, f) as $n \rightarrow \infty$.*

The intuition behind this definition is that the limit of this sequence of random variables should be thought of as a way to make sense of $(\Gamma_A, f1_{D \setminus A})$, which then has to be the same as (Γ_A, f) .

We leave it as an exercise to check that in the particular case of local sets where h_A is integrable, then this definition indeed coincides with the one given in the introduction. To do this, first one has to check that for all possible values of u of $D \setminus A_n$, $(\Gamma_A, f1_u)$ is a.s. equal to $\int_u h_A(z) f(z) dz$, we will come back to this in Section 5.

Finally, let us note already that the choice of working with dyadic approximations is somewhat arbitrary and the question whether changing this choice would change the definition is in fact not an easy one. Even if the examples that we will describe in the next section are clearly tailor-made for the particular definition, it is easy to adapt it to any other analogous choice. We will comment further on this in Section 5.

3. EXAMPLES OF “SMALL” NON-THIN LOCAL SETS.

In the present section, we prove the statements $(2)_d$: We construct and describe the main features of a particular local set of the d -dimensional GFF in $d \geq 2$, which is not thin, yet rather small.

3.1. An example using CLE_4 in two dimensions. Before we construct our actual examples, let us first very quickly describe how it is possible to use the coupling of the two-dimensional GFF with the Conformal Loop Ensembles CLE_4 to construct a local set which implies the statement $(2)_d$ when $d = 2$. Due to the fact that such a relationship is only known in dimension 2, this construction can not be generalized to higher dimensions, but it will nevertheless help understanding some features of the example that we will provide in the next subsections. Since this CLE_4 -based construction is not used in our main proofs, we choose here not to give a complete review of the Miller-Sheffield coupling of the CLE_4 with the GFF in two dimensions, and we refer the reader to [5] for background and details.

Let Γ be a GFF in a simply connected domain D . Recall that (see [5, 13]), it is possible to define deterministically from Γ a local set A_1 of Minkowski dimension $15/8$ such that the harmonic function h_{A_1} (that we denote by h_1) is constant and equal to $\pm 2\lambda$ in each connected component of $D \setminus A_1$, where here and throughout this section, 2λ is equal to the so-called height-gap $\sqrt{\pi}/2$ of the two-dimensional GFF. This set A_1 has the law of a CLE_4 , and the coupling just described is usually called the natural coupling of CLE_4 with the GFF.

Furthermore, as explained in [5], this local set is thin (in the present case, the definition of thin is the one given in the introduction because h_1 is integrable) and conditionally on A_1 , the sign of h_1 is chosen to be $+$ or $-$ independently in each connected component of $D \setminus A_1$.

We then define inductively, an increasing family A_n of local sets as follows: Suppose that for a given n , we have defined A_n in such a way that h_n is constant in each connected component of $D \setminus A_n$ and is equal to $2k\lambda$ for some integer $k \leq 1$. We then define A_{n+1} and h_{n+1} as follows:

- In the connected components of $D \setminus A_n$ where $h_n = 2\lambda$ we do nothing: these connected components are still in $D \setminus A_{n+1}$ and $h_{n+1} = 2\lambda$ there.
- In the other connected components, O , of $D \setminus A_n$, we construct, in O , the CLE_4 associated to the GFF Γ^{A_n} restricted to O . The connected components of $D \setminus A_{n+1} \cap O$ are defined to be the complement of this CLE_4 , and the values of the harmonic function are $h_{n+1} = h_n \pm 2\lambda$.

We finally define our local set A to be the closure of $\cup_n A_n$.

It is then easy that A is a local set, that h_A is equal to 2λ in each of the connected components of the complement of A . It is also easy to see that the Lebesgue measure of A is almost surely equal to 0, and we leave it as a simple exercise to the reader who has read [5] to check that in fact, it satisfies $(2)_2$.

Since $h_A = 2\lambda$, the set A can not be thin. Indeed, for any smooth non-negative test function f , the integral $\int_{D \setminus A} h_A(z) f(z) dz$ would be almost surely non-negative, and it can therefore not be the same random variable as $(\Gamma, f) - (\Gamma^A, f)$ (unless $f = 0$).

3.2. Another example in two dimensions. The previous example relies on the CLE technology which is not available for the GFF in higher dimensions. In the present subsection we first describe another local set A of the two-dimensional GFF that has a simple generalization when $d \geq 3$. One main feature is reminiscent of the previous case: We discover the GFF in a self-similar fashion (but we use the boundary of dyadic squares instead of nested CLE_4), and explore the GFF until its mean value in the dyadic square that we are currently looking at is likely to be positive, in some sense that we will make precise

Notation. Choose the domain D to be the unit square $(0, 1)^2$. As we are going to use nested dyadic squares, it is useful to use the following notation. We define S^\emptyset to be equal to D , and when u is a finite sequence of n elements of $\{1, \dots, 4\}$, then S^{u^1}, \dots, S^{u^4} are the four open dyadic subsquares of side-length 2^{-n-1} of S^u , where each one is a dyadic square of side-length 2^{-n} . We can for instance choose to associate the four indices respectively to the NW, NE, SW, SE subsquares. Thanks to this notation we can associate to each square a point in the tree $\{1, 2, 3, 4\}^*$, and a genealogy.

Let us also define for each dyadic square S^u , the random variable $\gamma_n(S^u) := (\Gamma_{T_n}, 1_{S^u})$, where T_n is the union of elements in \mathcal{T}_n . This is the conditional expectation of $(\Gamma, 1_{S^u})$ in S^u , when one observes the GFF outside (i.e. on the boundary) of the ancestor of S^u with height n if $n \leq |u|$ (the height of u), or the boundary of the childs of u with height n if $n > |u|$. It can also be viewed as (Γ, μ^u) where μ_n^u is a well-chosen measure supported on the boundary of the squares associated with S^u with height n .

We are going to discover progressively and simultaneously the GFF along the four segments from $(1/2, 0)$, $(1, 1/2)$, $(1/2, 1)$ and $(0, 1/2)$ to the middle point $(1/2, 1/2)$ (see the first image of figure Figure 1). When we have finished, then the unit square is divided into the four squares S^1, \dots, S^4 of side-length $1/2$. During this discovery, we can choose a modification of the conditional expectation of the random variable $(\Gamma, 1)$ (which is the mean value of Γ on S^\emptyset given the discovered values of the GFF in the four segments we have discovered) so that it evolves like a continuous martingale. Thus, we can parametrize time in such a way that at time t , we have discovered four segments of length $l(t)$ so that this conditional expectation has the law of a Brownian motion $B = B^\emptyset$ at time t .

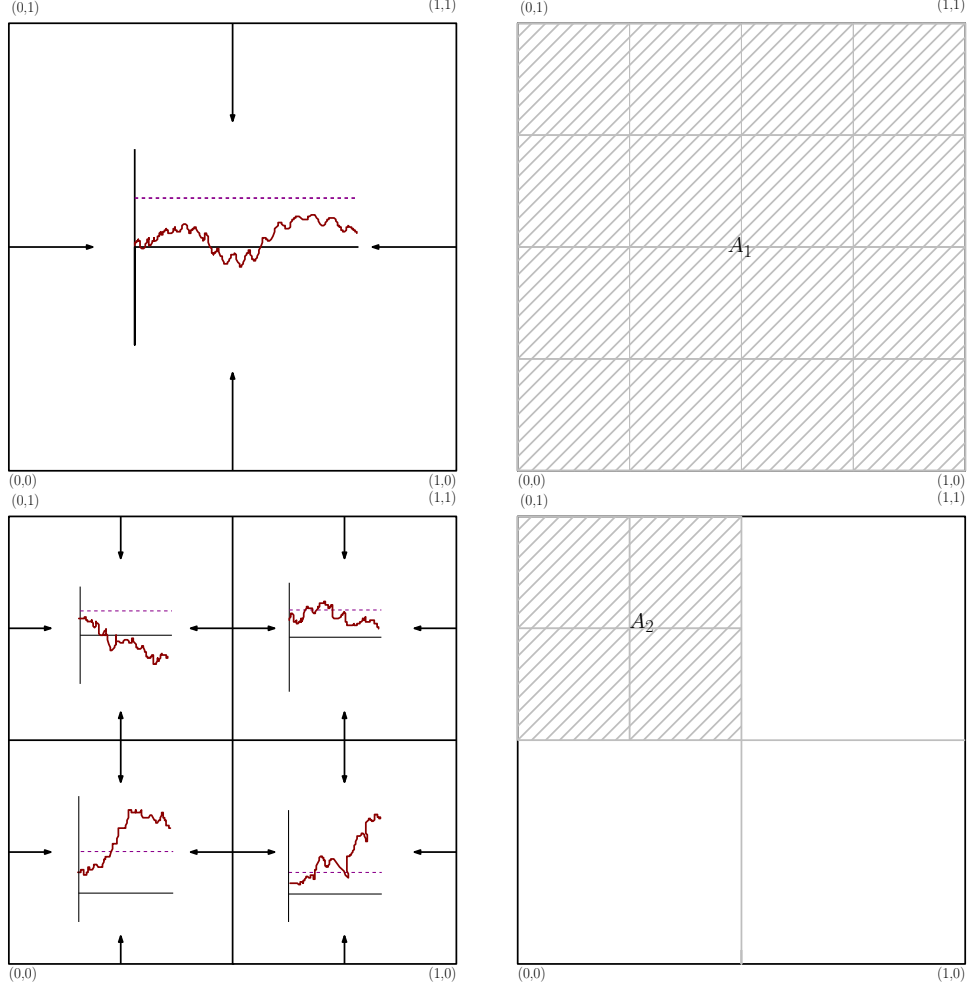


FIGURE 1. First two steps in the construction of A . In the left pictures we represent the Brownian motion associated to each point. In the right figure, the grey areas represents A_n .

Definition of A . If B hits 1 before time T , we define A^1 to be equal to the union of these four segments at the end-time T of this exploration, so that $U_1 := S^\emptyset \setminus A^1 = S^1 \cup \dots \cup S^4$. If not we take $A^1 = D$. Note that $\mathbb{E}[(\Gamma, 1) | \sup_{t \leq T} B \geq 1] = 1$.

If the Brownian motion has not reached 1 before time T , we continue exploring, and we do this independently and simultaneously in all four squares S^1, \dots, S^4 using the GFF Γ^{A^1} in each of them (note that Γ^{A^1} consists of four independent GFFs in the four squares). In each of these squares, we grow four boundary segments towards the center of the square, and we study the conditional expectation of $4(\Gamma^{A^1}, \mathbf{1}_{S^j})$ (the mean of the mass of Γ^{A^1} in S^j) given what one has discovered. By self-similarity, each of these four quantities evolve like four independent Brownian motions B^1, \dots, B^4 up to time T .

Now, in order to define A^2 , if $A^1 \neq D$, then $A^2 = A^1$, if not we look, for each S^i , at whether the BM $W^i := B(t \wedge T) + B^i(T - t)\mathbf{1}_{\{t \geq 0\}}$ hits level 1 before time $2T$ or not. A^2 is made by the closed union of all the squares of size 2^{-1} where this BM did not hit the level 1 before time $2T$, with the boundaries of all the squares of the same size where this event happen (see Figure 1). In other words, for each $n \geq 1$:

- The sets A^n and ∂A^n are local sets made out of the union of 2^{-n} dyadic segments with elements of \mathcal{S}_n , and A^n is such that $(A^n)_n = A^n$. We say that a square $s \in \mathcal{S}_n$ is still active (meaning that we will continue exploring inside it) when $s \in A^n$. Furthermore, active squares also come equipped with a Brownian motion W^s stopped at time Tn . We call K_n the set of active squares in \mathcal{S}_n and V_n the set of connected components of $D \setminus A^n$, i.e., the inactive components. Note that $V_n \subseteq \bigcup_{k=1}^n \mathcal{S}_k$.
- In order to construct A^{n+1} and to continue W , we proceed as follows: The components that were not active at step n remain inactive. For $s \in K_n$ continuously grow the middle lines as done in the first step and define for $0 \leq t \leq n(T+1)$ and s^+ any direct descendent of s , $W^{s^+}(t) := W^s(t \wedge nt) + B^s(t - nT)\mathbf{1}_{\{t \geq nT\}}$, where B^s is the BM associated to the change of the conditional expectation of $2^n(\Gamma^{A^n}, 1_s)$ given the increasing procedure in s . We keep active those squares s^+ where its associated BM did not hit 1 before time $(n+1)T$, and we make s^+ inactive (i.e. $s^+ \in V_m$ for $m \geq n+1$) if its associated BM hit 1 before time $(n+1)T$. We define A^{n+1} as the closed union of all the active squares at time $(n+1)$ with the boundary of the inactive squares. We can also see it as A^n minus the squares s^+ that became inactive in this step.

Note that A^n is non-increasing, and that the family V_n is non-decreasing. We define A to be the intersection of all A^n . The complement of A is then just the union of the squares that stop being active at some point, more precisely, $D \setminus A$ is the disjoint union of the squares in $\bigcup_n V_n$. Thus, we have that $A_n = A^n$. Note that for a given dyadic square s , on the even that $s \in V_n$, the harmonic function h_A coincides with the harmonic function $h_{D \setminus T_n}$ on s (where T_n the union of all boundaries of 2^{-n} -dyadic squares) and that $(\Gamma_A, 1_s) = \gamma_n(s)$.

The set A is not large. The construction shows immediatly that the probability that a given dyadic square s of side-length 2^{-n} is still active at step n is equal to the probability that a one-dimensional Brownian motion did not hit 1 before time $n \times T$, which decays like a constant times $1/\sqrt{n}$ as $n \rightarrow \infty$. From this, it follows readily that the size of A is indeed of the type required for $(2)_2$, i.e.:

Proposition 3.1. *The expected value of the area of the ε -neighborhood of A decays almost surely like $O(|\log \varepsilon|^{-1/2})$.*

Indeed, if $N_n = N_n(A)$ denotes the number of closed 2^{-n} dyadic squares that intersect A , then

$$\begin{aligned} \mathbb{E}[N_n] &= \sum_{s \in \mathcal{S}_n} \mathbb{E}[\mathbf{1}_{\{s \subseteq A_n\}}] + C \sum_{j=1}^{n-1} \sum_{s \in \mathcal{S}_j} 2^{n-j} \mathbb{E}[\mathbf{1}_{\{s \subseteq A_j \setminus A_{j+1}\}}] \\ &\leq 4^n \mathbb{P}(\text{BM does not hit one before } Tn) + C 2^n \sum_{j=1}^{n-1} j^{-3/2} 2^j \leq C \frac{4^n}{\sqrt{n}} \end{aligned}$$

(mind that in N_n , we have to also count the squares that intersect the boundaries of squares that have stopped being active, which explains the sum in j).

A first moment estimate. Note that in order to define the set A , we have in fact associated a Branching Brownian motion (BBM) W to each GFF, where each BM splits into 4 independently evolving BM at each time which is a multiple of T . However, it should be emphasized that for a given dyadic square s of side-length 2^{-n} , the value of the corresponding Brownian motion at time nT is not equal to the expected mean height of the GFF in s given the exploration up to the n -th generation. Indeed, this mean height has clearly a higher value when s is towards the centre of S than when it is near its boundary, which is not mirrored by the Branching Brownian motion

description. However, a key observation is that this difference is averaged out when one sums over all squares. For instance, it is easy to check by induction on n that

$$\sum_{s \in \mathcal{S}_n} \gamma_n(s) = \sum_{s \in \mathcal{S}_n} 4^{-n} W^s(nT),$$

if B^s denotes the Brownian motion that is following the branch of the BBM corresponding to s .

The variant of this result that will be useful for us is:

Lemma 3.2.

$$\mathbb{E}[(\Gamma_A, \mathbf{1}_{D \setminus A_n})] = \mathbb{E} \left[\sum_{s \in V_n} \text{Area}(s) \right].$$

The left-hand side is equal to the probability that a Brownian motion started from 0 does hit 1 before time nT , which converges to 1. This shows already that $(\Gamma_A, \mathbf{1}_{D \setminus A_n})$ can not converge in L^1 to $(\Gamma, 1)$, which is a symmetric random variable with mean 0.

Proof. Note that $D \setminus A_n = \bigcup_{s \in V_n} s$ and that at time n , $\Gamma_{\partial A_n} = \Gamma_A$ in all elements of V_n and $\Gamma_{\partial A_n} = \Gamma_{T_n}$ in all of those in K_n . This implies that $\mathbb{E}[(\Gamma_A, \mathbf{1}_{D \setminus A_n})] = -\mathbb{E}[\sum_{s \in K_n} \gamma_n(s)]$. Then, it is enough to prove that

$$\mathbb{E} \left[\sum_{s \in K_n} \gamma_n(s) \right] = \mathbb{E} \left[\sum_{s \in K_n} W^s \right] = -\mathbb{E} \left[\sum_{s \in V_n} \text{Area}(s) \right].$$

The second equality just follows from the stopping time theorem. For the first equality we have to work harder. Let us note that for all $s' \in \mathcal{S}_m$ and $s \in \mathcal{S}_n$ with ancestor s' , $W^s((m+1)T) - W^s(mT)$ is equal to $4^m(\gamma_{m+1}(s') - \gamma_m(s'))$ and that $\mathbb{E}[\mathbf{1}_{\{s \in K_n\}} \mid \Gamma_{T_{m+1}}]$ does not depend on s . Now, let us show that the increment of the harmonic function for $s \in K_n$ at level m can be computed using the Brownian motion,

$$\begin{aligned} \sum_{s \in \mathcal{S}^n: s' \leq s} \mathbb{E}[(\gamma_{m+1} - \gamma_m)(s) \mathbf{1}_{\{s \in K_n\}}] &= \sum_{s \in \mathcal{S}^n: s' \leq s} \mathbb{E}[(\gamma_{m+1} - \gamma_m)(s) \mathbb{E}[\mathbf{1}_{\{s \in K_n\}} \mid \Gamma_{T_{m+1}}]] \\ &= \sum_{s \in \mathcal{S}^n: s' \leq s} 4^{m-n} \mathbb{E}[(\gamma_{m+1} - \gamma_m)(s') \mathbb{E}[\mathbf{1}_{\{s \in K_n\}} \mid \Gamma_{T_{m+1}}]] \\ &= 4^{-n} \sum_{s \in \mathcal{S}^n: s' \leq s} \mathbb{E}[(W^s((m+1)T) - W^s(mT)) \mathbf{1}_{\{s \in K_n\}}]. \end{aligned}$$

We conclude by writing a $\sum_{s \in K_n} \gamma_n(s)$ as a telescopic sum. □

This set A is not thin. Our goal is now to derive the following fact, which combined with Proposition 3.1 proves the statement (2)₂:

Proposition 3.3. *This local set A is not thin.*

This is a direct consequence of the following claim:

Claim 3.4. *The sequence of random variables $(\Gamma_A, \mathbf{1}_{D \setminus A_n})$ is bounded in L^2 .*

Indeed, if $(\Gamma_A, \mathbf{1}_{D \setminus A_n})$ would converge in probability towards $(\Gamma, 1)$, then it would converge also in L^1 , and we have seen in the previous paragraph that this can not be the case.

Deriving Claim 3.4 requires some care. We have to bound covariances of the increments of the integral of the harmonic function in two squares, s and s' , at each step of the process. In order to do that, we separate the increments according to whether or not they come from the conditional expected value of T_m with m bigger or equal, p , the height of $s \wedge s'$, the last common ancestor of

s and s' . We realize that if we condition according to the values of the GFF in T_p there are many terms that become constant and allow us to go the increments of level p , instead of n .

In our proof, we use the following basic bound for centred Gaussian random variables X, Y : For any event E with positive probability,

$$(3.1) \quad \mathbb{E}[XY\mathbf{1}_E] \leq C \max\{\text{Var}(X), \text{Var}(Y)\} \mathbb{P}(E) \log(1/\mathbb{P}(E))$$

(to prove it, note first that due to the fact that $2ab \leq a^2 + b^2$, we can restrict ourselves to the case where $X = Y$, and by scaling it suffices to consider the case where X is a standard normal variable. The quantity $E[X^2\mathbf{1}_A]$ is maximal among all sets A with $P(A) = a$ for the set $A = \{X^2 > x\}$ where $P(X^2 > x) = a$, and the estimate then follows).

Proof of the claim. As in the beginning of Lemma 3.2, $(\Gamma_A, \mathbf{1}_{D \setminus A_n}) = (\Gamma_A, 1) + \sum_{s \in K_n} \gamma_n(s)$. Given that $\text{Var}(\Gamma_A, 1) \leq \text{Var}(\Gamma, 1)$ it is just enough to bound

$$\mathbb{E} \left[\sum_{s, s' \in K_n} \gamma_n(s) \gamma_n(s') \right].$$

We will do this by writing $\gamma_n(s)$ and $\gamma_n(s')$ as the sum of the increments at each iteration step. Things are a little bit messier than for the first moment, because one has more terms to evaluate. For $s, s' \in \mathcal{S}_n$, we will have to consider the common ancestor $w = s \wedge s'$. In the following lines, we first fix $p \geq 2$ and w a 2^{-p} -daydic square.

For any $m, o \geq p$ conditionally on Γ_{T_p} , $(\gamma_{m+1} - \gamma_m)(s) \mathbf{1}_{\{s \in K_n\}}$ and $(\gamma_{o+1} - \gamma_o)(s') \mathbf{1}_{\{s' \in K_n\}}$ are independent. Hence,

$$\begin{aligned} & \sum_{p \leq m, o < n} \sum_{\substack{s, s' \in \mathcal{S}^n \\ s \wedge s' = w}} \mathbb{E} [(\gamma_{m+1} - \gamma_m)(s) (\gamma_{o+1} - \gamma_o)(s') \mathbf{1}_{\{s, s' \in K_n\}}] \\ &= \sum_{p \leq m, o < n} \sum_{\substack{s, s' \in \mathcal{S}^n \\ s \wedge s' = w}} \mathbb{E} [\mathbb{E} [(\gamma_{m+1} - \gamma_m)(s) \mathbf{1}_{\{s \in K_n\}} \mid \Gamma_{T_p}] \mathbb{E} [(\gamma_{o+1} - \gamma_o)(s') \mathbf{1}_{\{s' \in K_n\}} \mid \Gamma_{T_p}]] \\ &= \sum_{\substack{s, s' \in \mathcal{S}^n \\ s \wedge s' = w}} 8^{-n} \mathbb{E} [\mathbb{E} [(W^s(nT) - W^s(pT)) \mathbf{1}_{\{s \in K_n\}} \mid \Gamma_{T_p}] \mathbb{E} [(W^{s'}(nt) - W^{s'}(pt)) \mathbf{1}_{\{s' \in K_n\}} \mid \Gamma_{T_p}]] \\ &\leq \sum_{\substack{s, s' \in \mathcal{S}^n \\ s \wedge s' = w}} 8^{-n} \mathbb{E} [(W^s(pT) + 1)^2 \mathbf{1}_{\{w \in K_p\}}] \leq C 8^{-p} \sqrt{p}, \end{aligned}$$

where for the third equality we used the same technique as in Lemma 3.2 and for the fourth and fifth we just use stopping time theorem for the BM B and for $B_t^2 - t$.

It is also true that $\mathbb{P}(u \in K_n \mid T_p)$ is constant for all u with ancestor w and that conditionally on Γ_{T_p} , $\{s \in K_n\}$ is independent of $\{s' \in K_n\}$. This allows us to compute the following second term

$$\begin{aligned} & \sum_{0 \leq m, o < p} \sum_{\substack{s, s' \in \mathcal{S}^n \\ s \wedge s' = w}} \mathbb{E} [(\gamma_{m+1} - \gamma_m)(s) (\gamma_{o+1} - \gamma_o)(s') \mathbf{1}_{\{s, s' \in K_n\}}] \\ &= \sum_{0 \leq m, o < p} \sum_{\substack{s, s' \in \mathcal{S}^n \\ s \wedge s' = w}} \mathbb{E} [(\gamma_{m+1} - \gamma_m)(s) \mathbb{P} [\mathbf{1}_{\{s \in K_n\}} \mid \Gamma_{T_p}] (\gamma_{o+1} - \gamma_o)(s') \mathbb{P} [\mathbf{1}_{\{s' \in K_n\}} \mid \Gamma_{T_p}]] \\ &= \mathbb{E} [\gamma_p(w)^2 \mathbf{1}_{\{w \in K_p\}}] \leq C 8^{-p} \sqrt{p} \log(p), \end{aligned}$$

where in the last step we have used (3.1) and the fact that the variance of $\gamma_p(w)$ is bounded by that of $(\Gamma, \mathbf{1}_w)$.

For the remaining term we need to bound the cross-product and using similar remarks as before we have that

$$\begin{aligned}
& \sum_{0 \leq m < p \leq o < n} \sum_{\substack{s, s' \in \mathcal{S}^n \\ s \wedge s' = w}} \mathbb{E} [(\gamma_{m+1} - \gamma_m)(s)(\gamma_{o+1} - \gamma_o)(s') \mathbf{1}_{\{s, s' \in K_n\}}] \\
&= \sum_{0 \leq m < p \leq o < n} \sum_{\substack{s, s' \in \mathcal{S}^n \\ s \wedge s' = w}} \mathbb{E} [(\gamma_{m+1} - \gamma_m)(s) \mathbb{P}[\mathbf{1}_{\{s \in K_n\}} \mid \Gamma_{T_p}] \mathbb{E}[(\gamma_{o+1} - \gamma_o)(s') \mathbf{1}_{\{s' \in K_n\}} \mid \Gamma_{T_p}]] \\
&= -\mathbb{E} [\gamma_p(w)(-W^w(pT) + 1)c(W^w(pT), n - p) \mathbf{1}_{\{w \in K_p\}}] \leq C 8^{-p} \sqrt{p} \log(p),
\end{aligned}$$

where $c(x, m)$ is the probability than a BM hits height $x + 1$ before time mT .

Summing all the previous terms up, we get that

$$\mathbb{E} \left[\sum_{s, s' \in K_n} \gamma_n(s) \gamma_n(s') \right] \leq C' + C \sum_{p=2}^{\infty} 4^{-p} \sqrt{p} \log(p) < \infty.$$

□

3.3. The example in higher dimensions. We now explain how to adapt the previous example to the higher-dimensional setting. The only slight difference is that in the two-dimensional case, we used the scale invariance of the GFF, while we will now use the scaling relation (2.1).

To adapt our example, let us define $D = S^\emptyset := (0, 1)^d$. We use the d -dimensional dyadic hypercubes denoted now by S^u where u are finite sequences in $\{1, \dots, 2^d\}$. When Γ is a GFF in D , we are now going to discover its values on all simultaneously growing all the $(d-1)$ -dimensional mid-hyperplanes. Then, the iterative construction proceeds in almost exactly the same way, but with a notable difference. Due to the different scaling behaviour of the GFF, if the evolution of the conditional mean height during the first iteration evolves like a Brownian motion up to some time T , then the evolution during the second iteration is that of a Brownian motion during time $T \times 2^{d-2}$, and so on. In other words, the intervals between the branching times of the branching Brownian motion will grow exponentially, and the n -th branching time will be $T_n = T(2^{(d-2)n} - 1)/(2^{d-2} - 1)$ instead of nT .

Other than that, nothing in the previous discussion changes. Lemma 3.2 together with Claim 3.4 become readily:

Lemma 3.5. *For this A we have that $\mathbb{E}[(\Gamma_A, \mathbf{1}_{D \setminus A_n})] = \mathbb{E}[\sum_{s \in V_n} \text{Volume}(s)]$ and the second moment of $(\Gamma_A, \mathbf{1}_{D \setminus A_n})$ is uniformly bounded.*

Just as in the 2-dimensional case, this then implies that A is not-thin.

To upper bound the Minkowski dimension, the only difference is that the probability that a given dyadic hypercube of side-length 2^{-n} is active at the n -th iteration is now the probability that a Brownian motion does not hit level 1 before time T_n , which leads to the estimate on the size of A as in $(2)_d$. Indeed, if N_n denotes the number of closed dyadic hypercubes that intersect A ,

$$\begin{aligned}
\mathbb{E}[N_n] &\leq C \sum_{j=1}^n \sum_{s \in \mathcal{S}^j} 2^{(n-j)(d-1)} \mathbb{E}[\mathbf{1}_{\{s \subseteq A_j\}}] = C 2^{n(d-1)} \sum_{j=1}^n 2^j \mathbb{P}(\text{BM hits 1 after time } T_n) \\
&\leq C 2^{n(d-1)} \sum_{j=1}^n 2^{(-d/2+2)j} \leq C 2^{\max\{d-1, d/2+1\}n}.
\end{aligned}$$

Thus, thanks the Markov inequality

$$\mathbb{P}[N_n \geq 2^{(\max\{d-1, d/2+1\}+\varepsilon)n}] \leq C 2^{-\varepsilon n},$$

and thanks to the Borel-Cantelli Lemma, we can conclude that the upper Minkowski dimension of A is almost surely bounded by $\max\{d-1, d/2+1\}$.

We conclude that $(2)_d$ holds for any $d \geq 3$.

Proposition 3.6 $((2)_d)$. *This local set A is not thin, and its upper Minkowski dimension is almost surely not larger than $\max\{d-1, d/2+1\}$.*

4. SMALL SETS ARE THIN (PROOF OF $(1)_d$)

Let us briefly note that the definition of thin sets can be extended to non-local sets: we say that a set A is thin if for all f smooth bounded function D we have that $(\Gamma, f\mathbf{1}_{A_n}) \xrightarrow{\mathbb{P}} 0$ as $n \rightarrow \infty$. This definition is useful because a.s.

$$\sum_{s \in \mathcal{S}_n: s \not\subseteq D \setminus A_n} (\Gamma, f\mathbf{1}_s) = (\Gamma, f\mathbf{1}_{A_n}),$$

so that it is sufficient to bound the value of the GFF in hyper-cubes of size 2^{-n} .

The following proposition links both definitions.

Lemma 4.1. *Let Γ be a GFF on D and A a local set. A is thin in this last sense if and only if A is a thin local set.*

Proof. It is enough to see that for all f smooth and bounded function :

$$(\Gamma, f\mathbf{1}_{A_n}) - ((\Gamma_A, f) - (\Gamma_A, f\mathbf{1}_{D \setminus A_n})) = (\Gamma^A, f\mathbf{1}_{A_n}) \xrightarrow{\mathbb{P}} 0 \quad \text{as } n \rightarrow \infty,$$

□

This shows for instance that any deterministic closed set A with zero Lebesgue measure is a thin local set. Indeed, if $\|f\|_\infty < 1$, by dominated convergence,

$$\mathbb{E} [(\Gamma, f\mathbf{1}_{A_n})^2] = \iint_{A_n \times A_n} f(x)G_D(x, y)f(y)dydx \rightarrow 0$$

as $n \rightarrow \infty$.

4.1. The case $d \geq 3$. The idea of the proof will be just to get uniform bounds on the mean values of Γ on elements of \mathcal{S}_n by second moment estimates and Borel-Cantelli arguments, and to then use Lemma 4.1 to conclude that if our sets are small enough then they are thin.

Let us recall that there exists an absolute constant C_d such that for any $s \in \mathcal{S}_n$ and any bounded function f ,

$$(4.1) \quad \iint_{s \times s} f(x)G(x, y)f(y)dx dy \leq C_d 2^{-(d+2)n},$$

which readily implies the following:

Lemma 4.2. *Let $d \geq 3$, $D \subseteq \mathbb{R}^d$ be an open set and Γ a GFF in D . For any $\beta < d/2 + 1$ and any smooth bounded function f , almost surely,*

$$\lim_{n \rightarrow \infty} \sup_{s \in \mathcal{S}_n} |(\Gamma, f\mathbf{1}_s)| 2^{\beta n} = 0.$$

Proof. For $n \in \mathbb{N}$ and $s \in \mathcal{S}_n$, $(\Gamma, f\mathbf{1}_s)$ is a centered Gaussian random variable with variance $\iint_{s \times s} f(x)G_D(x, y)f(y)dx dy$, so that by (4.1),

$$\sum_{n \in \mathbb{N}} \sum_{s \in \mathcal{S}_n} \mathbb{P}(|(\Gamma, f\mathbf{1}_s)| \geq \varepsilon 2^{-n\beta}) < \infty,$$

and we can conclude using the Borel-Cantelli Lemma. □

This lemma now enables to prove the statement $(1)_d$ for $d \geq 3$:

Proposition 4.3 $((1)_d)$. *Let $D \subseteq \mathbb{R}^d$ be an open set, Γ a GFF in D and A a local set of Γ . If the upper Minkowski dimension of A is almost surely strictly smaller than $\max\{d-2, d/2+1\}$, then A is a thin local set.*

Proof. Let us first note that if the upper Minkowski dimension $\delta(A)$ of A is strictly smaller than $d-2$, then A is polar, so that Lemma 2.1 implies that $\Gamma_A = 0$, and that A is thin local set.

Let us now assume instead that $\delta(A) < d/2 + 1$ almost surely. The following argument will in fact not use the fact that A is a local set: Let us fix $f \in C_0^\infty(\mathbb{R}^d)$ define the following events:

$$\tilde{\Omega}_\beta := \left\{ N_n = o(2^{n(d/2+1-\beta)}) \text{ as } n \rightarrow \infty \right\}.$$

Since $\delta(A) < d/2 + 1$, we know that $\cup_{\beta>0} \tilde{\Omega}_\beta$ holds almost surely. On the other hand, for each given β , if $\tilde{\Omega}_\beta$ holds, then by Lemma 4.2, we see that $(\Gamma, f\mathbf{1}_{A_n})$ tends to 0. Hence, we conclude that this convergence holds in fact almost surely and by Lemma 4.1 we conclude that A is thin. \square

Note that with this proposition and its proof we can get some other basic properties of thin sets.

Corollary 4.4. *Let $D \subseteq \mathbb{R}^d$ be an open set, Γ a GFF on D and A, B thin local sets. If the upper Minkowski dimension of A is strictly smaller than $d/2 + 1$, then:*

- (1) $A \cup B$ is also a thin set.
- (2) If A is a local set such that h_A is integrable (i.e., such that $\int_{D \setminus A} |h_A| < \infty$) and B has zero Lebesgue measure, then a.s. $B \setminus A$ is thin for $\Gamma^A := \Gamma - \Gamma_A$.

Proof. (1) Note that for any bounded smooth function f :

$$|(\Gamma, \mathbf{1}_{(A \cup B)_n})| \leq |(\Gamma, f\mathbf{1}_{B_n})| + |(\Gamma, f\mathbf{1}_{A_n \setminus B_n})| \xrightarrow{\mathbb{P}} 0, \quad \text{as } n \rightarrow \infty,$$

where the second term goes to 0 because it can be written as a sum over elements of \mathcal{S}_n and the amount of terms in that sum is smaller than the cardinal $\{s \in \mathcal{S}_n : s \subseteq A_n\}$, so the same argument used in the last proof to show that A is thin can be applied.

- (2) Let f be a bounded function and note that the fact that because h_A is integrable and B has 0 measure $\int_{D \setminus A} h_A(x) \mathbf{1}_{(B \setminus A)_n}(x) dx$ goes to 0. Additionally $(\Gamma, f\mathbf{1}_{(B \setminus A)_n})$ because of the same reason as in the proof of the last fact. \square

In future work we plan to prove that when the upper Minkowski dimension of A is smaller than $d/2 + 1$, then h_A integrable on $D \setminus A$, which will allow to relax a little bit the conditions in this last corollary, see [4].

Note that this does not answer the question whether the fact that B is thin implies that its Lebesgue measure is 0. Remark that such statements are non-trivial, due for instance to the fact that we cannot exclude at this point, the very unlikely fact that there exist thin local sets, with non-thin subsets.

4.2. The case $d = 2$. A useful fact in two dimensions is that in order to prove that a set is thin, it is not necessary to test against all smooth bounded functions:

Lemma 4.5. *Let A be a local set of a GFF Γ in a bounded domain $D \subseteq \mathbb{R}^2$. Then A is thin as soon as $\sum_{s \in \mathcal{S}_n : s \not\subseteq D \setminus A_n} |(\Gamma, \mathbf{1}_s)| \rightarrow 0$ in probability as $n \rightarrow \infty$.*

Proof. Let $s \in \mathcal{S}_n$ and $f \in C_0^\infty(\mathbb{R}^2)$. Define $\bar{f}_s := \int_s f(x) dx / \text{Leb}(s)$ as the mean of f in s . We have that

$$\mathbb{E} \left[(\bar{f}_s(\Gamma, \mathbf{1}_s) - (\Gamma, \mathbf{1}_s f))^2 \right] \leq C \|f\|_{C_0^1} n 2^{-6n}.$$

Using Borel Cantelli we have that the $\sup_{s \in \mathcal{S}_n} (\Gamma, \mathbf{1}_s(f - \bar{f}_s))$ is $O(n^2 2^{-3n})$. Due to the fact that the cardinal of \mathcal{S}_n is of the order 4^n ,

$$\sup_{B \subseteq \Gamma(D)} \left| \sum_{s \in \mathcal{S}_n : s \notin D \setminus A_n} \bar{f}_s(\Gamma, \mathbf{1}_s) - (\Gamma, \mathbf{1}_s f) \right| \leq \sum_{s \in \mathcal{S}_n : s \notin D \setminus A_n} |(\Gamma, \mathbf{1}_s(f - \bar{f}_s))| \xrightarrow{a.s.} 0 \quad \text{as } n \rightarrow \infty.$$

Thus, it is enough to show the convergence to 0 of $|\sum_{s \in \mathcal{S}_n : s \notin D \setminus A_n} \bar{f}_s(\Gamma, \mathbf{1}_s)|$, which clearly follows from the hypothesis. \square

Now, we prove (1)₂, i.e., the following statement:

Proposition 4.6 ((1)₂). *Let $D \subseteq \mathbb{C}$ be a bounded open set, Γ a GFF on D and A a random closed set. If there exists $\delta > 0$ such that*

$$\mathbb{E}[N_n] = o(4^n/n^{1/2+\delta})$$

as $n \rightarrow \infty$, then A is a thin set.

Proof. Thanks to the scaling properties it is enough to work with a domain $D \subseteq B(0, 1/4)$. Take Γ a GFF in D and $s \in \mathcal{S}_n$. Let us define B_s the event where $|(\Gamma, \mathbf{1}_s)|$ is bigger than the quantity

$$m_n := 4^{-n} \sqrt{n} \sqrt{\log n + 2 \log \log n} \sqrt{\log 2} / \sqrt{2\pi},$$

and R_n the number of squares where this inequality holds. We also define

$$S_n = \sum_{s \in \mathcal{S}_n} |(\Gamma, \mathbf{1}_s)| \mathbf{1}_{B_s}.$$

Note S_n converges to 0 in L^1 . Indeed, $\text{Var}(\Gamma, \mathbf{1}_s) \leq 4^{-n} \sqrt{n} \sqrt{\log(2)} / \sqrt{2\pi}$ and:

$$\mathbb{E}[S_n] = \sum_{s \in \mathcal{S}_n} \mathbb{E}[|(\Gamma, \mathbf{1}_s)| \mathbf{1}_{B_s}] \leq C \sqrt{n} \exp\left(-\frac{\log(n) + 2 \log \log(n)}{2}\right) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Now, let us note that if $|\{s \in \mathcal{S}_n : s \notin D \setminus A_n\}| \leq R_n$, then $\sum_{s \in \mathcal{S}_n : s \notin D \setminus A_n} |(\Gamma, \mathbf{1}_s)| \leq S_n$. Thus, we have that for all $\eta > 0$

$$\begin{aligned} & \mathbb{P}\left(\sum_{s \in \mathcal{S}_n, s \notin D \setminus A_n} |(\Gamma, \mathbf{1}_s)| \geq \eta\right) \\ & \leq \mathbb{P}\left(|\{s \in \mathcal{S}_n : s \notin D \setminus A_n\}| \geq \frac{4^n}{n^{1/2+\delta}}\right) + \mathbb{P}(S_n \geq \eta) + \mathbb{P}\left(R_n \leq \frac{4^n}{n^{1/2+\delta}}\right). \end{aligned}$$

where the first term goes to 0 thanks to Markov inequality, and the second thanks to the fact that S_n converges to 0 in L^1 . Thanks to Lemma 4.5, the only thing we need to finish the proof is to show that the last term converges to 0. In order to do that let us first prove that $\mathbb{E}(R_n) \geq C 2^{2n} n^{-1/2-\delta/2}$ and $\text{Var}(R_n) = O(2^{4n} n^{-4/3})$.

For $x, y \in D \setminus T_n$ take s_x the only element in \mathcal{S}_n where it belongs and define

$$\begin{aligned} \alpha_x &:= \frac{m_n}{\sqrt{\text{Var}((\Gamma, \mathbf{1}_{s_x}))}} \\ u_{x,y} &:= \frac{\text{Cov}((\Gamma, \mathbf{1}_{s_x}), (\Gamma, \mathbf{1}_{s_y}))}{\sqrt{\text{Var}((\Gamma, \mathbf{1}_{s_x}))} \sqrt{\text{Var}((\Gamma, \mathbf{1}_{s_y}))}}. \end{aligned}$$

We note that for all $(x, y) \in E^{(n)} := \{(x, y) \in D^2 : \min(|x - y|, d(y, \partial D), d(x, \partial D)) \geq 1/n\}$

$$\begin{aligned}\alpha_x &\geq (1 - C \log(n)/n) \times \sqrt{\log n + 2 \log \log(n)}, \\ u_{x,y} &\leq C \frac{\log(1/|x - y|)}{n}.\end{aligned}$$

Using the lower bound for α_x and basic Gaussian estimates we can compute the lower bound for the first moment

$$\mathbb{E}[R_n] \geq C 4^n (\log(n))^{-1/2} \exp\left(-\frac{\log n + \log \log N}{2}\right) \geq C \frac{4^n}{n^{1/2+\delta/2}}$$

The variance estimate is more involved and we use the lower bound for the correlation. First let us note that

$$4^{-2n} \text{Var}[R_n] = \iint_{D \times D} [\mathbb{P}(\bar{\Gamma}_x \geq \alpha_x, \bar{\Gamma}_y \geq \alpha_y) - \mathbb{P}(\bar{\Gamma}_x \geq \alpha_x) \mathbb{P}(\bar{\Gamma}_y \geq \alpha_y)] dx dy,$$

where $\bar{\Gamma}_x = (\Gamma, \mathbf{1}_{s_x})/\text{Var}(\Gamma, \mathbf{1}_{s_x})$, and $(\bar{\Gamma}_x, \bar{\Gamma}_y)$ has the law of the centred Gaussian vector whose coordinates have variance 1 and the covariance between them is $u_{x,y} \geq 0$. Note the fact that $\mathbb{P}(\bar{\Gamma}_x \geq \alpha_x) = O(n^{-1/2})$ and that the volume of $D \times D \setminus E^{(n)}$ is $O(n^{-1})$ implies that we only need to bound the term for $x, y \in E^{(n)}$. In this case we just note that by definition of Gaussian vector the term inside the integral is equal to (we denote $u_{x,y} = u$)

$$\begin{aligned}& \frac{1}{2\pi} \int_{\alpha_x}^{\infty} \int_{\alpha_y}^{\infty} e^{-\frac{1}{2}(a^2+b^2)} \left(\frac{1}{\sqrt{1-u^2}} \exp\left(-\frac{u^2 a^2 + u^2 b^2 - 2uab}{2(1-u^2)}\right) - 1 \right) db da \\ & \leq \frac{1}{2\pi} \int_{\alpha_x}^{\infty} \int_{\alpha_y}^{\infty} e^{-\frac{1}{2}(a^2+b^2)} \left(\frac{1}{\sqrt{1-u^2}} - 1 \right) da db - \frac{C \log(|x - y|)}{n} \int_{\alpha_x}^{\infty} \int_{\alpha_y}^{\infty} a b e^{-(a^2+b^2-4uab)/2} db da\end{aligned}$$

where to get the inequality we have used that for $x \geq 0$, $\exp(x) - 1 \leq x \exp(x)$. A Taylor expansion shows that the first term is an $O(n^{-2})$. Note that the second term is smaller than a constant times $n^{-1} \log(1/|x - y|) \mathbb{E}[XY \mathbf{1}_{\{X \geq \alpha_x, Y \geq \alpha_y\}}]$, where (X, Y) is a Gaussian vector such that each term has variance $1/(1 - 4u^2)$; using that $XY \leq X^2 + Y^2$ it is easy to see that this term is an $O(n^{-4/3})$. Integrating over $(x, y) \in E^{(n)}$ we get the desired bound on the variance.

To conclude we note that $\mathbb{E}[R_n] \geq 4^n/n^{1/2+\delta}$ and

$$\mathbb{P}\left(R_n \leq \frac{4^n}{n^{1/2+\delta}}\right) \leq \mathbb{P}\left((\mathbb{E}[R_n] - R_n)^2 \geq \left(\mathbb{E}[R_n] - \frac{4^n}{n^{1/2+\delta}}\right)^2\right) = \text{Var}[R_n] \left(\mathbb{E}[R_n] - \frac{4^n}{n^{1/2+\delta}}\right)^{-2};$$

the estimates on the mean and the variance of R_n show that this term is an $O(n^{-1/3+\delta})$ which concludes the proof. \square

5. SOME COMMENTS ABOUT THE DEFINITIONS OF THIN LOCAL SETS

Let us now make some somewhat abstract comments about the definition of local sets. One general strategy in order to define local sets is to use some deterministic “enlargements” of the random sets A (see for instance [28]). To the best of our knowledge, only dyadic-type enlargements have been used in earlier work, but this is a rather arbitrary choice. For our purposes here, it seems natural to consider also other possible deterministic enlargements – indeed, this a priori choice could be important, given that some property may hold for one approximation scheme, and not for the other.

Let us describe one possible class of discrete approximation schemes (DAS), for which the proofs of the present paper can be adapted rather directly.

DAS when $d \geq 3$. Define a pre-DAS for a domain $D \subseteq \mathbb{R}^d$ to be a sequence $(\mathcal{A}_n)_{n \geq 0}$ of families of closed sets $\mathcal{A}_n = (\mathcal{B}_n, \mathcal{C}_n)$ for which there exists some (large) constant $C \in \mathbb{R}$ such that the following holds for any $n \in \mathbb{N}$:

- (1) For any two distinct c and c' in \mathcal{C}_n , the Lebesgue measure of $c \cap c'$ is zero.
- (2) For any c in \mathcal{C}_n the diameter of c is upper bounded by $C2^{-n}$ and its volume is lower bounded by $2^{-nd}/C$.
- (3) $\text{Leb}(\bigcup_{b \in \mathcal{B}_n} b) = 0$. And for all $E \subseteq \mathbb{R}^d$ compact, the cardinal of the elements of \mathcal{B}_n that intersect E is finite

For a fixed pre-DAS \mathcal{A}_n , take $\mathcal{B}^n := \bigcup_{b \in \mathcal{B}_n} b$, the set of all points covered by elements of \mathcal{B} . For all closed set $A \subseteq \bar{D}$, define $\mathcal{A}\{A\}_n$ as the set of all elements of \mathcal{C}_n that have a non empty intersection with $A \setminus \mathcal{B}^n$ and take $\mathcal{A}[A]_n$ the union of all sets in $\mathcal{A}\{A\}_n$ with all the set in \mathcal{B}_n that have non-empty intersection with A . More formally,

$$\begin{aligned} \mathcal{A}\{A\}_n &:= \{c \in \mathcal{C}_n : c \cap A \setminus \mathcal{B}^n \neq \emptyset\}, \\ \mathcal{A}[A]_n &:= \bigcup_{c \in \mathcal{A}\{A\}_n} c \cup \bigcup_{\substack{b \in \mathcal{B}_n, \\ b \cap A \neq \emptyset}} b. \end{aligned}$$

We then say that a pre-DAS \mathcal{A}_n is a DAS if for all closed set $A \subseteq \bar{D}$, $\mathcal{A}[A]_n \searrow A$.

In this context we understand $\mathcal{A}[A]_n$ as an approximation of A using a union of elements in \mathcal{B}_n and \mathcal{C}_n , it should be understood that the elements of \mathcal{C}_n are the only ones “giving mass” to $\mathcal{A}[A]_n$. $\mathcal{A}\{A\}_n$ represents all the set in \mathcal{C}_n that where used to construct $\mathcal{A}[A]_n$.

Dyadic hyper-cubes provide an example of DAS – more precisely, when \mathcal{C}_n are the closed dyadic hypercubes of side-length 2^{-n} intersected with D and \mathcal{B}_n is empty. This is our canonical DAS and it is such that for all closed sets A the cardinal of $\mathcal{A}\{A\}_n$ is N_n .

Note that condition (2) implies that if A is bounded $|\mathcal{A}\{A\}_n| \leq CN_n$ and that there exists an absolute constant C_d such that for any $c \in \mathcal{C}_n$

$$(5.1) \quad \iint_{c \times c} f(x)G(x, y)f(y)dxdy \leq C_d 2^{-(d+2)n}.$$

DAS when $d = 2$. In two dimensions, we will modify slightly the definitions. A pre-DAS for a domain $D \subseteq \mathbb{R}^2$ is now a countable collection of families of closed sets $\mathcal{A}_n = (\mathcal{B}_n, \mathcal{C}_n)$ for which (1) and (3) holds and (2) is replaced another condition (2') that we now describe.

For each $c \in \mathcal{C}_n$, let us define

$$G_D(c_x) = \iint_{c_x \times c_x} G_D(x, y)dxdy, \quad G_D(c_x, c_y) = \iint_{c_x \times c_y} G_D(x, y)dxdy.$$

We will say that (2') holds if for all c and c' in \mathcal{C}_n that are at distance greater than $1/n$ from each other and from ∂D ,

$$(5.2) \quad 1 - \frac{C \log n}{n} \leq \frac{1}{\sqrt{2\pi}} \frac{2^{-2n} \sqrt{n}}{\sqrt{G_D(c_x)}} \leq 1 + \frac{C \log n}{n},$$

$$(5.3) \quad u_{x,y} := \frac{G_D(c_x, c_y)}{G_D(c_x)G_D(c_y)} \leq C \frac{-\log(|x - y|) + 1}{n},$$

Note that this is the type of estimates that hold in the case of dyadic squares, and that we have used in our arguments.

We are now ready to give an alternative definition of thin local sets. This definition coincides with that of [28] in the special case where h_A is integrable on $D \setminus A$ (so that working with DAS is not necessary). It is also similar to Lemma 3.10 [14], where they ask Γ to be a.s. determined by the restriction of Γ to $D \setminus A$. On the other hand the first example presented in Section 3 is non-thin,

but it is proven in [3] that Γ is a function of the restriction of Γ to $D \setminus A$ (because A is measurable with respect to this restriction).

Definition 5.1. Let Γ be a Gaussian Free Field on a domain D and $A \subseteq D$ a local set. We say that A is a thin set if for all f smooth and with bounded support in \mathbb{R}^d ($C_0^\infty(\mathbb{R}^d)$) and for all DAS \mathcal{A}_n , the sequence $(\Gamma_A, f\mathbf{1}_{D \setminus \mathcal{A}[A]_n})$ converges in probability to (Γ_A, f) when $n \rightarrow \infty$.

Note that $(\Gamma_A, f\mathbf{1}_{D \setminus \mathcal{A}[A]_n})$ is always well defined thanks to the fact that when the $\text{supp}(f)$ is compact, $\mathcal{A}[A]_n$ can take only finitely many values. Also, as we have said before, if $\int_{D \setminus A} |h_A| < \infty$, then the limit of $(\Gamma_A, f\mathbf{1}_{D \setminus \mathcal{A}[A]_n})$ is a.s. equal to $\int_{D \setminus A} f(z)h_A(z)$ and this limit does not depend on the chosen DAS. The DAS framework is relevant in the case where the integral of $|h_A|$ on $D \setminus A$ diverges.

Additionally, when $\int_{D \setminus A} |h_A| < \infty$ it is actually enough to check the criteria for functions f in $C_0^\infty(D)$, because when we approximate one function in $C_0^\infty(\mathbb{R}^d)$ restricted to D by one in $C_0^\infty(D)$ both the left and right term of the definition converge to what they should.

Let us briefly note that the definition of thin sets can be extended to non-local sets: We say that a set A is thin if for all $f \in C_0^\infty(\mathbb{R}^d)$ and for all DAS \mathcal{A} .

$$\sum_{c \in \mathcal{A}\{A\}_n} (\Gamma, f\mathbf{1}_c) = (\Gamma, f\mathbf{1}_{\mathcal{A}[A]_n}) \xrightarrow{\mathbb{P}} 0, \quad \text{as } n \rightarrow \infty.$$

it is easy to see that Lemma 4.1 also holds in this setup. This, together with the estimates (5.1), (5.2), and (5.3), allows us to prove that when a deterministic set A has 0 Lebesgue measure it is thin and the equivalents of Proposition 4.3 and 4.6. Note that (2') is necessary if we want to adapt the proof of Proposition 4.6. Note that if a set is a thin local set, then it is thin under the definition of Section 2.3. This implies that the sets A defined in Section 3 are not thin local sets.

Let us conclude with the following general remarks: It is an open question whether thinness for one approximation scheme implies thinness for all of them. Another issue is that it does not allow to capture the fact that we are also asking our sets to be also local; our proofs that rely on the approximation schemes do not only to prove local thinness but also thinness. This is related to the fact that so far, the question whether the union of thin local sets is always a local set is also still open.

REFERENCES

- [1] J. Aru. The geometry of the Gaussian free field combined with SLE processes and the KPZ relation, PhD thesis, 2015
- [2] J. Aru. KPZ relation does not hold for the level lines and SLE_κ flow lines of the Gaussian free field *Probab. Theory Related Fields* 163(3-4): 465–526, 2015.
- [3] J. Aru, T. Lupu, A. Sepúlveda. First passage sets of the 2D continuum GFF. In preparation.
- [4] J. Aru, A. Sepúlveda. Survey in preparation.
- [5] J. Aru, A. Sepúlveda, W. Werner. On bounded-type thin local sets of the two-dimensional Gaussian free field, preprint, 2016.
- [6] E. Bolthausen, J.D. Deuschel, G. Giacomin. Entropic repulsion and the maximum of the two-dimensional harmonic crystal, *Ann. Probab.*, 29(4): 1670–1692, 2001.
- [7] J. Dubédat. SLE and the free field: partition functions and couplings. *Journal of the American Mathematical Society*, 22(4): 995–1054, 2009.
- [8] B. Duplantier, S. Sheffield. Liouville quantum gravity and KPZ, *Invent. Math.*, 2: 333–393, 2011.
- [9] J.B. Garnett and D. E. Marshall. Harmonic Measure, *Cambridge University Press*, 2005.
- [10] X. Hu and J. Miller and Y. Peres. Thick points of the Gaussian Free Field, *Ann. Probab.* 38(2): 896–926, 2010.
- [11] Y. Le Jan. Markov Path, Loops and Fields L.N. in Math 2026, *Springer*, 2011.
- [12] T. Lupu. From loop clusters and random interlacement to the free field, *Ann. Probab.*, to appear.
- [13] J. Miller and S. Sheffield. The GFF and $\text{CLE}(4)$, slides of 2011 talks and private communications.
- [14] J. Miller and S. Sheffield. Imaginary Geometry I. Interacting SLEs, *Probab. Theor. rel. Fields*, to appear.

- [15] J. Miller and S. Sheffield. Quantum Loewner Evolution, *Duke Math. J.* 165(17): 3241–3378 (2016).
- [16] J. Miller and S. Sheffield. Liouville quantum gravity and the Brownian map I: The QLE $(8/3, 0)$ metric, preprint 2015.
- [17] J. Miller, S.S. Watson and D.B. Wilson. Extreme nesting in the conformal loop ensemble, *Ann. Probab.*, 44(2), 1013–1052, 2016.
- [18] P. Mörters and Y. Peres. Brownian motion, Cambridge University Press, 2010.
- [19] E. Nelson. Construction of quantum fields from Markoff fields, *J. Funct. Anal.*, 12, 97 (1973)
- [20] Ş. Nacu and W. Werner. Random soups, carpets and fractal dimensions. *J. London Math. Society*, 83, 789–809, 2011.
- [21] Yu. A. Rozanov. Markov Random Fields. *Springer-Verlag* 1982.
- [22] O. Schramm. Scaling limits of loop-erased random walks and uniform spanning trees. *Israel J. Math.* 118: 221–288, 2000.
- [23] O. Schramm and S. Sheffield. A contour line of the continuum Gaussian free field. *Probability Theory Related Fields*, 157: 47–80, 2013.
- [24] O. Schramm, S. Sheffield, and D. B. Wilson. Conformal radii for conformal loop ensembles. *Communications in Mathematical Physics*, 288(1): 43–53, 2009.
- [25] W. Qian, W. Werner. Decomposition of Brownian loop-soup clusters. preprint 2015.
- [26] S. Sheffield. Gaussian free fields for mathematicians. *Probab. Theory Related Fields*, 139(3–4): 521–541, 2007.
- [27] S. Sheffield and W. Werner. Conformal Loop Ensembles: The Markovian characterization and the loop-soup construction. *Ann. Math.*, 176, 1827–1917, 2012.
- [28] W. Werner. Topics on the GFF, Lecture Notes, 2016.

DEPARTMENT OF MATHEMATICS, ETH ZÜRICH, RÄMISTR. 101, 8092 ZÜRICH, SWITZERLAND
E-mail address: leonardo.sepulveda@math.ethz.ch